Appendix for "The Effect of Large-Scale Anti-Contagion Policies on the Coronavirus (COVID-19) Pandemic"

Solomon Hsiang,*[1,2] Daniel Allen,[†1] Sébastien Annan-Phan,[†1,3] Kendon Bell,[†1,4] Ian Bolliger,[†1,5] Trinetta Chong,[†1] Hannah Druckenmiller,[†1,3] Andrew Hultgren,[†1,3] Luna Yue Huang,[†1,3] Emma Krasovich,[†1] Peiley Lau,[†1,3] Jaecheol Lee,[†1,3] Esther Rolf,[†1,6] Jeanette Tseng,[†1] and Tiffany Wu[†1]

[1]Global Policy Laboratory, Goldman School of Public Policy, UC Berkeley
[2]National Bureau of Economic Research & Centre for Economic Policy Research
[3]Agricultural & Resource Economics, UC Berkeley
[4]Manaaki Whenua, Landcare Research
[5]Energy & Resources Group, UC Berkeley
[6]Electrical Engineering & Computer Science Department, UC Berkeley

*Correspondence: shsiang@berkeley.edu. [†] Equal contribution (alphabetical).

The following sections provide additional information to support the article and to document our data collection and processing procedures.
All data and code are available at https://github.com/bolliger32/gpl-covid.
Additional resources and updates are available at www.globalpolicy.science/covid19.
For repository related questions, please contact bolliger@berkeley.edu.
For data related questions, please contact peiley@berkeley.edu.

This document last updated: March 21, 2020. Data last updated: March 19, 2020.

1. Appendix Tables and Figures

Table A1: Number of policies tabulated by administrative divisions of each country.

Country	Adm0	Adm1	Adm2	Adm3	Total
China	-	4 Province	126 City	-	130 All
France	8 Country	2 Région	49 Département	-	59 All
Iran	5 Country	16 Province	-	-	21 All
Italy	12 Country	24 Region	81 Province	77 Municipality	194 All
South Korea	20 Country	22 Province	-	-	42 All
United States	16 Country	191 State	38 County	245 City	490 All
Total	**61**	**259**	**294**	**322**	**936**

Table A1: Policy data have been collected at various levels of administrative divisions in each country. “Adm0” represents the country level, and higher “Adm*” numbers indicate smaller administrative subdivisions. Each policy is counted at the highest level of specificity of the regions where the policy is applied. For example, if a country has ten regions at the “Adm1” level, and a policy is applied across five of those regions, the policy is counted as five separate “Adm1” policies rather than a single “Adm0” policy.

Figure A1: Error distributions for estimated growth rates of COVID-19 cases by country.

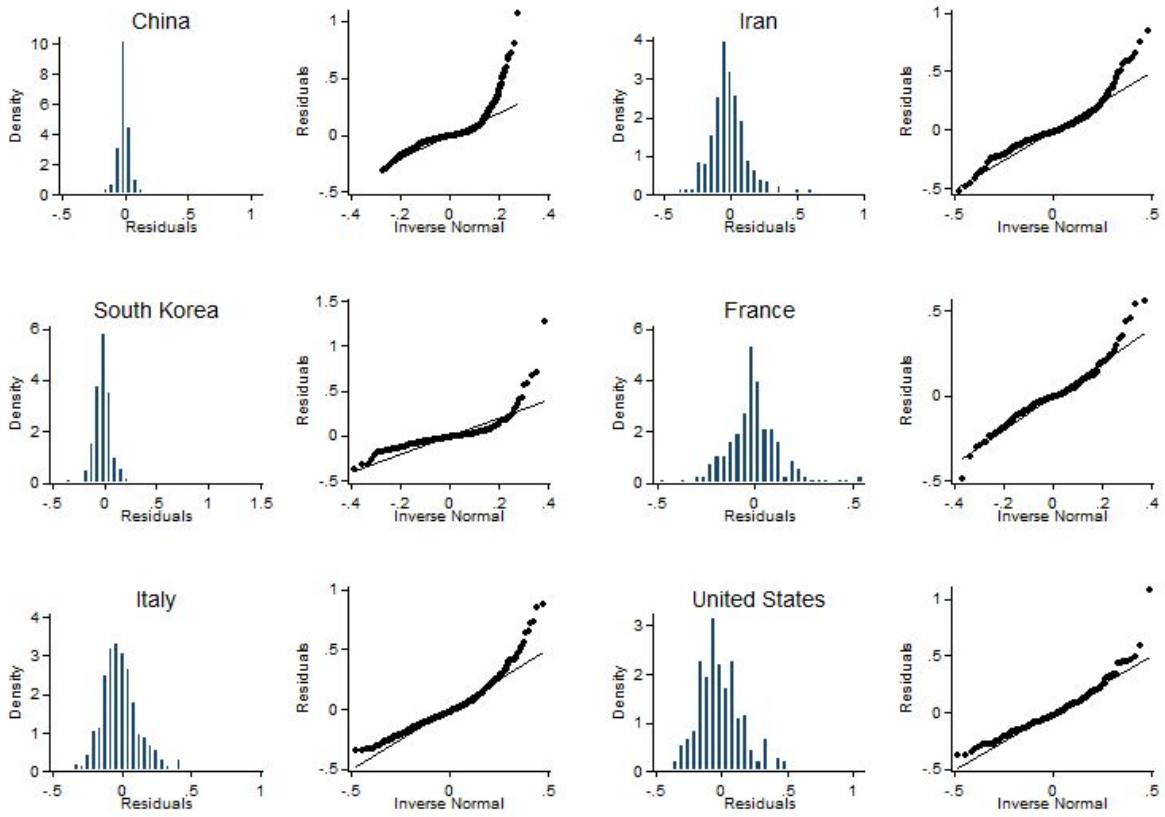
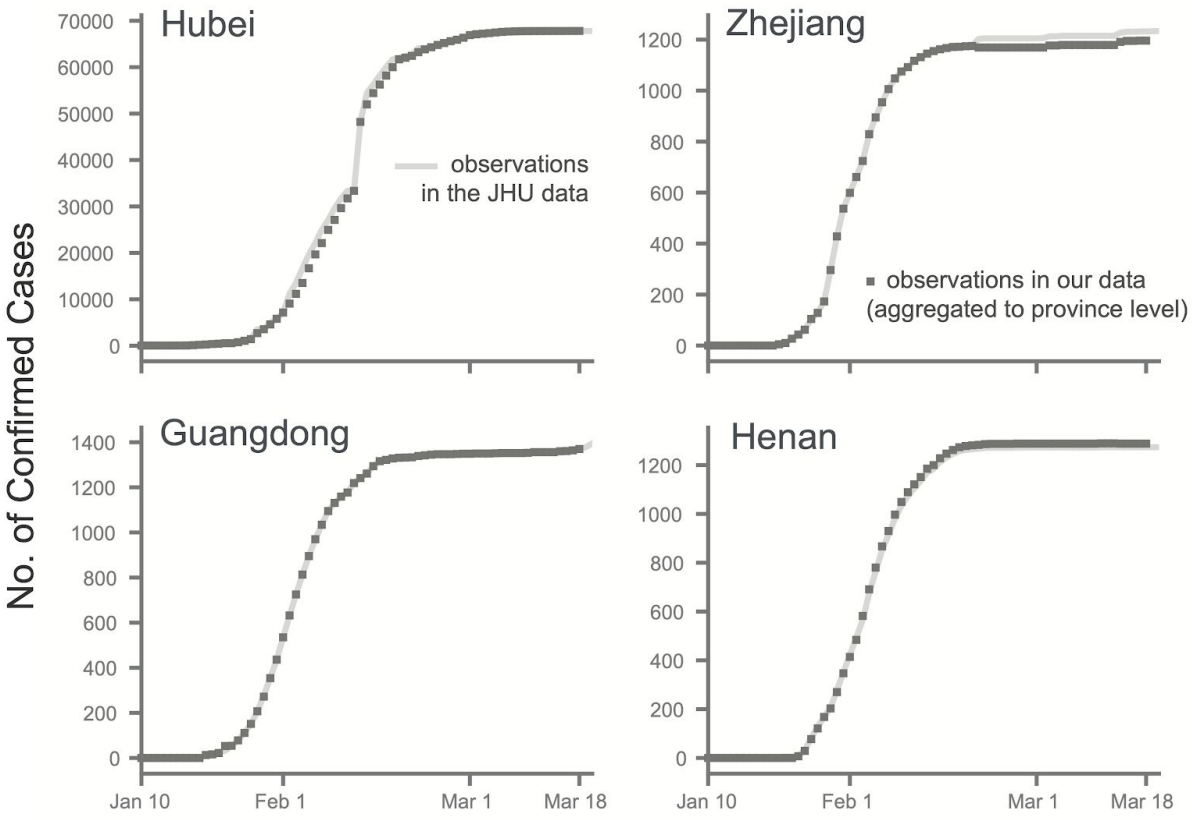


Figure A1: These plots show the error structure for each country-specific econometric model used to predict the daily growth of active or cumulative COVID-19 cases under the country's actual policy regime, as compared to the counterfactual world where no policies were enacted. See the full model under the Methods - Econometric analysis section as well as the results in Figure 3 of the main paper.

Figure A2: Validating our disaggregated epidemiological data against data from the Johns Hopkins Center for Systems Science and Engineering.

Panel A: The four provinces in China with the highest number of cumulative confirmed cases as of March 18, 2020.



Panel B: The entire country of South Korea.

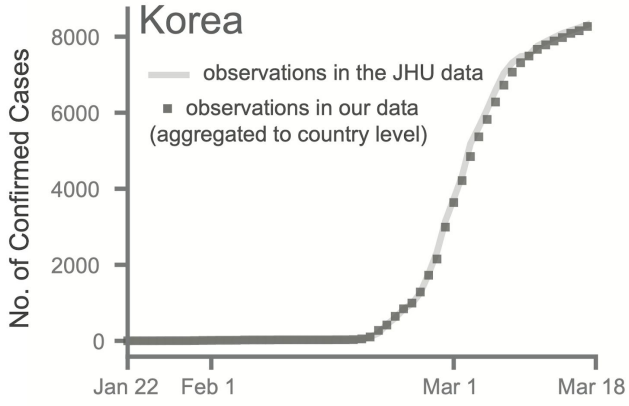


Figure A2: As an additional check, we compared the cumulative number of confirmed cases from a handful of regions in our collated epidemiological dataset to the same statistics from the 2019 Novel Coronavirus COVID-19 (2019-nCoV) Data Repository by the Johns Hopkins Center for Systems Science and Engineering (JHU CSSE).[1] We conducted this comparison for the two countries that we had the most data for and at two different administrative levels. In China, we aggregate city level data up to the province level, and in Korea we aggregate provincial level data up to the country level. The numbers tracked each other for the entire time series we have collected thus far.

2. Data Acquisition and Processing

This section describes the data acquisition and processing procedure for both epidemiological and policy data used in this paper. The sources for both types of data come from a variety of in-country data sources, which include government public health websites, regional newspaper articles, and Wikipedia crowd-sourced information. We have supplemented this data with international data compilations. A list of the epidemiological and policy data compiled for this analysis can be found here.

Epidemiological Data

The epidemiological datasets and sources used in this paper are described below. The main health variables of interest:

1. "*cum_confirmed_cases*": The total number of confirmed positive cases in the administrative area since the first confirmed case.
2. "*cum_deaths*": The total number of individuals that have died from COVID-19.
3. "*cum_recoveries*: The total number of individuals that have recovered from COVID-19.
4. "*cum_hospitalized*": The total number of hospitalized individuals.
5. "*cum_hospitalized_symptom*": The total number of symptomatic hospitalized individuals.
6. "*cum_intensive_care"* : The total number of individuals that have received intensive care.
7. "*cum_home_confinement*": The total number of individuals that have been self-quarantined in their homes as a result of a positive test.
8. "*active_cases*": The number of individuals who currently still test positive on the date of the observation.
9. "*active_cases_new*": The number of new cases since the previous date.
10. "*cum_tests*": The total number of tests (includes both positive and negative results) conducted in an administrative unit.

Additional metadata accompanying the health outcome variables:

1. "*date*": The date of observation.
2. "*adm0_name*": The ISO3 code to which this observation belongs.
3. "*adm1_name*": The name of the "Adm1" region to which this observation belongs.

[1] https://github.com/CSSEGISandData/COVID-19

4. "*adm2_name*": If the dataset contains observations at the "Adm2" level, then this is the name of the "Adm2" region to which this observation belongs.
5. "*adm[1,2]_id*": Any alphanumeric ID scheme to identify different administrative units (e.g. FIPS code).
6. "*lat*": The latitude of the centroid of the administrative unit.
7. "*lon*": The longitude of the centroid of the administrative unit.
8. "*policies_enacted*": The number of active policies that are in place for the administrative unit as of that date. This variable is not population weighted.
9. "*testing_regime*": A categorical variable used to identify when an administrative region (or country) changed their COVID-19 testing regime. This is zero-indexed, with the ordering only indicating chronological progression (there is no external meaning to Regime 2 vs. Regime 1 vs. Regime 0, and there is no consistency enforced for coding across countries). For example, if China changes their testing regime twice, all observations prior to the first regime change would be coded "*testing_regime*=0," all observations in between the two changes would be coded "*testing_regime*=1," and all observations after the second change would be coded "*testing_regime*=2."

Data Imputation:

In instances where health outcome observations are missing or suffer from data quality issues, we have imputed to fill in the missing values. Imputed health outcome variables are denoted by "*[health_outcome]_imputed.*" For the majority of our analyses we do not use imputed data; France is the exception where we impute two days of missing data. We do this to ensure we have variation in policy variables for use in the analysis.

We impute by:

1. Taking the natural log of the non-missing observations pertaining to that health outcome variable.
2. Linearly interpolating over the missing dates for that health outcome variable.
3. Exponentiating the interpolated values back into levels and rounding to the nearest integer.

China

We have collated a city level time series health outcome dataset in China for 339 cities from January 10, 2020 to present-day.

For data from January 24, 2020 onwards, we relied on the public dataset Ding Xiang Yuan[2] (DXY) that reports daily statistics across Chinese cities. Since DXY only publishes the most recent (cross-sectional) statistics (and not the historical data), we used the time series dataset scraped from DXY in an open source GitHub project[3]. The web scraper program checks for updates at least once a

[2] https://ncov.dxy.cn/ncovh5/view/pneumonia
[3] https://github.com/BlankerL/DXY-COVID-19-Data

day for the statistics published on DXY and records any changes in the number of cumulative confirmed cases, cumulative recoveries or cumulative deaths.

We assumed that no updates to the statistics meant there had been no new cases. We dropped a small number of cases that had been recorded but not assigned to a specific city (many of these cases are imported ones from other cities). We also dropped confirmed cases in prison populations (we assumed the spread of COVID-19 in prisons was not affected by the implementation of city-level lockdowns or travel ban policies).

For city level health outcomes prior to January 24, 2020, we manually collected official daily statistics from the central[4] and provincial (Hubei,[5] Guangdong,[6] and Zhejiang[7]) Chinese government websites. We did not collect city level health outcomes recorded prior to January 24, 2020 in provinces that had fewer than ten confirmed cases at that date. We made this decision since our analysis dropped observations with fewer than ten cumulative confirmed cases to prevent noisy data during the early transmission phase from disproportionately biasing the estimated results.

After merging the two datasets, we conducted a few quality checks:

(1) We checked that cumulative confirmed cases, cumulative recoveries, and cumulative deaths were increasing over time. In instances when cumulative outcomes decreased over time, we assumed that the recent numbers were more reliable, and treated the earlier number of cumulative cases as missing (this was often due to data entry errors or cases where patients that were reported to have been diagnosed with COVID-19, but were later found out to actually have tested negative). The magnitude of these errors was relatively small. We filled in any missing data with the imputation methodology described in the health data overview section.

(2) We validated our city level dataset by aggregating observations up to the provincial level and comparing the time trends from the aggregated dataset to that of the provincial dataset collated by Johns Hopkins University.[8] We confirmed that the two datasets matched very closely (see Figure A2 Panel A).

Testing Regime Changes:
As of the time of writing, the criteria for being diagnosed with COVID-19 had changed twice in China.[9] On February 13, 2020, China recategorized patients who exhibited symptoms, as determined through a chest scan, as part of the "confirmed" cases count even if they had not tested positive in the PCR test. This was due to concerns that the PCR test had relatively high false negative rates. On February

[4] http://www.nhc.gov.cn/xcs/yqtb/list_gzbd.shtml
[5] http://wjw.hubei.gov.cn/bmdt/ztzl/fkxxgzbdgrfyyq/xxfb/index_26.shtml
[6] http://wsjkw.gd.gov.cn/zwyw_yqxx/index_5.html
[7] http://www.zj.gov.cn/col/col1228996608/index.html
[8] https://github.com/CSSEGISandData/COVID-19
[9] https://www.cnbc.com/2020/02/26/confusion-breeds-distrust-china-keeps-changing-how-it-counts-coronavirus-cases.html

20, 2020, China reversed this decision. We included this information in the dataset because it could have potentially changed the levels and short-term growth rates of the number of confirmed cases.

France

We have collated a regional level time series health outcome dataset in France from February 15, 2020 to present-day.

We used the number of confirmed COVID-19 cases by *région* from France's government website.[10] The sources listed for this dataset were the French public health website,[11] the Ministry of Solidarity and Health,[12] French newspapers that reported government information,[13] and regional public health websites.[14] Given that this dataset was not published on a daily basis, we supplemented it by scraping the number of confirmed cases by *région* on the French public health website, which has been updated every day.[15]

Testing Regime Changes:
As of the time of writing, there have been no changes in France's testing regime.

South Korea

We have collated a provincial level time series health outcome dataset in South Korea from January 20, 2020 to March 14, 2020.

Most provinces in South Korea have been publishing data on their number of confirmed coronavirus cases. Daegu,[16] Gyeongsangbuk-do,[17] Jeollabuk-do,[18] Gyeongsangbuk-do,[19] and Sejong[20] provinces

---

[10] https://www.data.gouv.fr/en/datasets/fr-sars-cov-2/

[11] https://www.santepubliquefrance.fr/maladies-et-traumatismes/maladies-et-infections-respiratoires/infection-a-coronavirus/

[12] https://solidarites-sante.gouv.fr/soins-et-maladies/maladies/maladies-infectieuses/coronavirus/article/points-de-situation-coronavirus-covid-19

[13] https://france3-regions.francetvinfo.fr/

[14] https://www.ars.sante.fr/

[15] https://www.santepubliquefrance.fr/maladies-et-traumatismes/maladies-et-infections-respiratoires/infection-a-coronavirus/articles/infection-au-nouveau-coronavirus-sars-cov-2-covid-19-france-et-monde

[16] http://www.daegu.go.kr/dgcontent/index.do?menu_id=00936642&menu_link=/icms/bbs/selectBoardArticle.do&bbsId=BBS_02112

[17] https://www.gb.go.kr/Main/open_contents/section/wel/page.do?mnu_uid=5857&LARGE_CODE=360&MEDIUM_CODE=90&SMALL_CODE=10mnu_order=2

[18] http://www.jeonbuk.go.kr/board/list.jeonbuk?boardId=BBS_0000105&menuCd=DOM_000000105010004000&contentsSid=1189&cpath=

[19] https://www.gb.go.kr/

have been reporting the number of confirmed cases on a daily basis. For these provinces, we recorded this published health data.

Given that the province of Gangwon-do[21] does not report provincial level health data, we refer to the daily number of new cases reported by each of its counties (Chuncheon-si,[22] Wonju-si,[23] Gangneung-si,[24] Taebaek-si,[25] Sokcho-si,[26] and Samcheok-si[27]). As a result, we manually collected the number of new confirmed cases from each county's webpage and aggregated the numbers to the provincial level.

The remaining provinces (Seoul,[28] Gyeonggi-do,[29] Incheon,[30] Busan,[31] Ulsan,[32] Gwangju,[33] Chungcheongnam-do,[34] Chungcheongbuk-do,[35] Gyeongsangnam-do,[36] Jeju,[37] and Jeollanam-do[38]) did not explicitly publish the number of cumulative confirmed cases. However, they did publish patient-level data, including the date of when patients had tested positive. For these provinces, we constructed the measure of cumulative confirmed cases by counting the number of daily confirmed cases and adding it to the previous date's total.

Most provinces did not publish the number of deaths. Instead, we checked the daily policy briefings posted on the government homepages mentioned in the footnotes and manually collected mortality data. In instances when mortality data were not found in the briefings, we obtained the mortality data from other official sources, such as through social media sources (e.g. Facebook) and blogs maintained by local governments. Lastly we supplement these sources with mortality data reported in news articles.

Testing regime changes:
We collected information on testing regime changes from the homepage of the Korean Center for Disease Control and Prevention (KCDC). In the press release menu, the KCDC uploaded daily briefing announcements which contained information on testing criteria and changes to the testing regime.[39]

---

[20] https://www.sejong.go.kr/bbs/R3273/list.do?cmsNoStr=17465
[21] https://www.provin.gangwon.kr/covid-19.html
[22] https://www.chuncheon.go.kr/index.chuncheon?menuCd=DOM_000000599001000000
[23] https://www.wonju.go.kr/intro.jsp
[24] https://www.gn.go.kr/
[25] http://www.taebaek.go.kr/intro.jsp
[26] http://www.sokcho.go.kr/intro.html
[27] http://www.samcheok.go.kr/02179/02696.web
[28] http://www.seoul.go.kr/coronaV/coronaStatus.do
[29] https://www.gg.go.kr/contents/contents.do?ciIdx=1150&menuId=2909
[30] https://www.incheon.go.kr/health/HE020409
[31] http://www.busan.go.kr/corona19/index
[32] http://www.ulsan.go.kr/corona.jsp
[33] https://www.gwangju.go.kr/
[34] http://www.chungnam.go.kr/coronaStatus.do
[35] http://www1.chungbuk.go.kr/covid-19/index.do
[36] http://xn--19-q81ii1knc140d892b.kr/main/main.do
[37] https://www.jeju.go.kr/corona19.jsp
[38] https://www.jeonnam.go.kr/coronaMainPage.do
[39] https://www.cdc.go.kr/board/board.es?mid=a30402000000&bid=0030

Initially, the South Korean government only tested people who: 1) demonstrated respiratory symptoms within 14 days after visiting Wuhan South China Seafood Wholesale Market and 2) those who had pneumonia symptoms within 14 days after returning from Wuhan.[40]

As the outbreak spread, the KCDC broadened the criteria for testing. Starting January 28, 2020, the agency isolated 1) those who had fever or respiratory symptoms upon returning from Hubei province and 2) those who had symptoms of pneumonia upon returning from mainland China.[41,42] We coded this as the first change in the testing regime.

The second testing regime change occurred on February 4, 2020, when the KCDC announced that people who had had any "routine contacts" with confirmed cases were required to self quarantine for a 14-day period. The agency defines two categories of contacts: close contacts and routine contacts. The former is defined as a person who has been within two meters of, in the same room as, or exposed to any respiratory secretions of an infected individual. The latter refers to whether the individual conducted any activity in the same place and time as the infected person. Prior to this regime change, KCDC separated those two cases and applied different quarantine policies; starting February 4, 2020, any routine contacts were also required to be self-quarantined. [43]

Shortly thereafter, South Korea aggressively expanded the scope of their testing. Starting February 7, 2020, the KCDC broadened the definition of suspected cases to 1) anyone who developed a fever or respiratory symptoms within 14 days after returning from China, 2) anyone who developed a fever or respiratory symptoms within 14 days after being in close contact with a confirmed case, and 3) anyone suspected of contracting COVID-19 based on their travel history to affected countries and their clinical symptoms.[44] Moreover,the KCDC announced that the test would be free for all suspected cases and

[40] https://www.cdc.go.kr/board/board.es?mid=a20501000000&bid=0015&list_no=365654&act=view
[41] https://www.cdc.go.kr/board/board.es?mid=a20501000000&bid=0015&list_no=365874&act=view
[42] NB: The KCDC English website explains the testing regime change in a more condensed format: "Any citizens identified with a fever or respiratory symptoms and have visited Wuhan will be isolated and tested at a nationally designated isolation hospital, and any foreigners staying in Korea will be conducted in cooperation with police."
https://www.cdc.go.kr/board/board.es?mid=a30402000000&bid=0030&act=view&list_no=365888&tag=&nPage=1
[43]
http://www.mohw.go.kr/react/al/sal0301vw.jsp?PAR_MENU_ID=04&MENU_ID=0403&page=1&CONT_SEQ=352662
[44]
https://www.cdc.go.kr/board/board.es?mid=a30402000000&bid=0030&act=view&list_no=366114&tag=&nPage=1
NB: The date of this press release is February 8, 2020, but the definition of "suspected cases" was effective starting from February 7, 2020.

confirmed cases.[45] As a result of these efforts, KCDC announced that they would begin to test 3,000 people daily, a marked increase from only 200 people a day.[46]

The KCDC revised their guidelines on February 20, 2020 in order to test more people. Their press release stated: "Suspected cases with a medical professional's recommendation, regardless of travel history, will get tested. Additionally, those who are hospitalized with unknown pneumonia will also be tested. Lastly, anybody in contact with a diagnosed individual will need to self-isolate, and will only be released when they test negative on the thirteenth day of isolation."[47]

As the number of patients grew rapidly, the KCDC decided to focus on more vulnerable groups. In their February 29, 2020 press release, the agency stated: "The KCDC has asked local government and health facilities to focus on tests and treatment, especially targeting those aged 65+ and those with underlying conditions who need early detection and treatment." This change was coded as our last testing regime change in the dataset.[48]

Italy

We have collated a regional and provincial level time series health outcome dataset in Italy from February 24, 2020 to present-day.

This data came from the GitHub repository maintained by the Italian Department of Civil Protection (*Dipartimento della Protezione Civile*). Health outcomes included the number of confirmed cases, the number of deaths, the number of recoveries, and the number of active cases. These figures have been updating daily at 5 or 6 pm (Central European Time). The regional level dataset was pulled directly from "*dati-regioni/dpc-covid19-ita-regioni.csv,*" and the provincial level dataset was pulled from "*dati-province/dpc-covid19-ita-province.csv.*"

Testing regime changes:
The testing regime change in Italy occurred when the Director of Higher Health Council announced on February 26, 2020 that COVID-19 testing would only be performed on symptomatic patients, as the majority of the previous tests performed were negative.

[45] NB: The testing fee was already somewhat affordable; a person needed to pay 160,000 KRW (about $130 USD). A related article can be found here: https://www.edaily.co.kr/news/read?newsId=02604326625668224&mediaCodeNo=257

[46] http://www.mohw.go.kr/upload/viewer/skin/doc.html?fn=1581054767217_20200207145247.pdf&rs=/upload/viewer/result/202003/

[47] https://www.cdc.go.kr/board/board.es?mid=a30402000000&bid=0030&act=view&list_no=366247&tag=&nPage=4#

[48] https://www.cdc.go.kr/board/board.es?mid=a30402000000&bid=0030&act=view&list_no=366406&tag=&nPage=2

Iran

We have collated a provincial level time series health outcome dataset in Iran from February 19, 2020 to present-day.

The Iranian government had been announcing its new daily number of COVID-19 confirmed cases at the provincial level on the Ministry of Health's website. This data has been compiled daily in the table "New COVID-19 cases in Iran by province"[49] located in the "2020 coronavirus pandemic in Iran" article on Wikipedia.

We spot-checked the data in the Wikipedia table against the Iranian Ministry of Health announcements[50] using a combination of Google Translate and a comparison[51] of the numbers in the announcements (which were written in Persian script) to the Persian numbers.

Testing regime changes:
On March 6, 2020, the Ministry of Health announced[52] a national coronavirus plan, which included contacting families by phone to identify potential cases, along with the disinfecting of public places. The plan was to begin in the provinces of Qom, Gilan, and Isfahan, and then would be rolled out nationwide. On March 13, 2020, the government announced a military-enforced home isolation policy throughout the nation.[53] This announcement included nationwide disinfecting of public places. While a follow-up announcement of the March 6 high testing regime stating its complete rollout was not found, the March 13 announcement did reference the implementation of the public spaces component of the earlier plan across the country. We thus assumed that the high testing regime had also been fully rolled out on March 13, 2020.

United States

We have collated a state level time series health outcome dataset in the United States from January 22, 2020 to present-day.

The data comes from the Github repository associated with the Johns Hopkins University (JHU) interactive dashboard (Dong, Du & Gardner 2020, Lancet). As of the time of writing, the data are available here. The repository and dashboard are updated essentially in real-time; at least daily.

Testing regime changes:

---

[49] https://en.wikipedia.org/wiki/2020_coronavirus_pandemic_in_Iran
[50] Example of Ministry of Health data
http://behdasht.gov.ir/index.jsp?siteid=1&fkeyid=&siteid=1&pageid=54782&newsview=200716
[51] Google Translate sometimes translates various Persian numbers as "1". Persian numbers compared here: https://www.languagesandnumbers.com/how-to-count-in-persian/en/fas/
[52]
https://www.dailymail.co.uk/news/article-8082443/ANOTHER-senior-Iranian-official-dies-coronavirus.html
[53] https://www.theguardian.com/world/2020/mar/13/revolutionary-guards-enforce-coronavirus-controls-iran

To determine the testing regime, we used estimated daily counts of the cumulative number of tests conducted in every state, as aggregated by the largely crowdsourced effort named "The Covid Tracking Project" (covidtracking.com). We estimated the total number of tests as the sum of confirmed positive and negative cases. For some states and some days, there have been no negative case counts, in which case we utilize just the confirmed positive cases. We also ensured that the confirmed number of positive cases agreed with the counts in the JHU dataset.

We programmatically filtered for possible testing regime changes by filtering for any consecutive days during which the testing rate increased at least 250% from one day to the next, and where this jump was an increase of at least 150 total tests over one day. After visually inspecting the candidates, we removed detected testing regime changes for North Carolina and Connecticut, as these states did not demonstrate spikes in their testing rate, but rather a more gradual and steady rate in the increase of testing.

(NB: the last download from covidtracking.com was March 19, 19:30 PST. We have been updating the process and the removal of detected testing regime changes periodically, so this may change.)

Policy Data

The policy events, datasets, and sources used in this paper are described below. For each country, the relevant country-specific policies identified were then mapped to a harmonized policy categorization used across all countries.

The policy categories are coded as binary variables, where "*[policy_variable]*" = 0 before the policy is implemented in that area, and "*[policy_variable]*" = 1 on the date the policy is implemented (and for all subsequent dates until the policy is lifted). The main policy categories identified across the six different countries fall into four broad classes:

1. Restricting travel:
    a. "*travel_ban_local*" : A policy that restricts people from entering or exiting the administrative area (e.g county or province) treated by the policy.
    b. "*travel_ban_intl_in*": A policy that either bans foreigners from specific countries from entering the country, or requires travelers coming from abroad to self-isolate upon entering the country.
    c. "*travel_ban_intl_out*": A policy that suspends international travel to specific foreign countries that have high levels of COVID-19 outbreak.
    d. "*travel_ban_country_list*": A list of countries for which the national government has issued a travel ban or advisory. This information supplements the policy variable "*travel_ban_intl_out.*"
    e. "*transit_suspension*": A policy that suspends any non-essential land-, rail-, or water-based passenger or freight transit.

2. Distancing through cancellation of events and suspension of educational/commercial/religious activities:
    a. "*school_closure*": A policy that closes school and other educational services in that area.
    b. "*business_closure*": A policy that closes all offices, non-essential businesses, and non-essential commercial activities in that area. "Non-essential" services are defined by area.
    c. "*religious_closure"*: A policy that prohibits gatherings at a place of worship, specifically targeting locations that are epicenters of COVID-19 outbreak. See the section on Korean policy for more information on this policy variable.
    d. "*work_from_home*": A policy that requires people to work remotely. This policy may also include encouraging workers to take holiday/paid time off.
    e. "*event_cancel*": A policy that cancels a specific pre-scheduled large event (e.g. parade, sporting event, etc). This is different from prohibiting all events over a certain size.
    f. "*no_gathering*": A policy that prohibits any type of public or private gathering. (whether cultural, sporting, recreational, or religious). Depending on the country, the policy can prohibit a gathering above a certain size, in which case the number of people is specified by the "*No_gathering_size*" variable.
    g. "*no_demonstration*": A policy that prohibits protest-specific gatherings. See the section on Korean policy for more information on this policy variable.
    h. "*social_distance*": A policy that encourages people to maintain a safety distance (often between one to two meters) from others. This policy differs by country, but includes other policies that close cultural institutions (e.g. museums or libraries), or encourage establishments to reduce density, such as limiting restaurant hours.
3. Quarantine and lockdown:
    a. "*pos_cases_quarantine*": A policy that mandates that people who have tested positive for COVID-19, or subject to quarantine measures, have to confine themselves at home. The policy can also include encouraging people who have fevers or respiratory symptoms to stay at home, regardless of whether they tested positive or not.
    b. "*home_isolation*": A policy that prohibits people from leaving their home regardless of their testing status. For some countries, the policy can also include the case when people have to stay at home, but are allowed to leave for work- or health-related purposes. For the latter case, when the policy is moderate, this is coded as '*home_isolation* = 0.5.'
4. Additional policies
    a. "*emergency_declaration*": A decision made at the city/municipality, county, state/provincial, or federal level to declare a state of emergency. This allows the affected area to marshal emergency funds and resources as well as activate emergency legislation.
    b. "*paid_sick_leave*": A policy where employees receive pay while they are not working due to the illness.

Optional policies:

In the cases when the aforementioned policies are optional, we denote this as "*[policy_variable]_opt.*"

Population weighting of policy variables:

In the cases when only a portion of the administrative unit (e.g. half of the counties within the state) are affected by the implementation of the policy, we weight the policy variable by the percentage of population within the administrative unit that is treated by the policy. This is denoted as "*[policy_variable]_popwt,*" and the value that this variable can take on is a continuous number between 0 and 1. Sources for the population data are detailed in a later section.

China

We obtain data on China's policy response to the COVID-19 pandemic by culling data on the start dates of travel bans and lockdowns at the city-level from the "2020 Hubei lockdowns" Wikipedia page, [54] the Wuhan Coronavirus Timeline project on Github,[55] and various news reports.

To combat the spread of COVID-19, the Chinese government imposed travel restrictions and quarantine measures, starting with the lockdown of the city of Wuhan, the origin of the pandemic, on January 23, 2020. Immediately following the Wuhan lockdown, neighboring cities followed suit, banning travel into and out of their borders, shutting down businesses, and placing residents under household quarantine. The same policy measures were implemented in cities across China for the next three weeks.

Some lockdowns occurred during the national Chinese New Year holiday (January 24 - 30, 2020) when schools and most workers were on break. On January 27, 2020, China extended the official holiday to February 2, 2020, while many additional provinces delayed resuming work and opening schools for even longer.[56] The Chinese New Year holiday is analogous to containment policies such as school closures and restrictions on non-essential work. We do not specifically estimate the effect of this holiday extension, as most cities were in lockdown during the extended holiday, and a lockdown is a more restrictive containment measure. A lockdown requires all residents to stay home, except for medical reasons or essential work, and only allows one person from each household to go outside once every one to five days (exact policy varied by city).

France

We obtain data on France's policy response to the COVID-19 pandemic from the French government website, press releases from each regional public health site, and Wikipedia.

---

[54] https://en.wikipedia.org/wiki/2020_Hubei_lockdowns
[55] https://github.com/Pratitya/wuhan2020-timeline
[56] https://www.china-briefing.com/news/china-extends-lunar-new-year-holiday-february-2-shanghai-february-9-contain-coronavirus-outbreak/

The French government website contains a timeline of all national policy measures.[57] Each regional public health agency (*l'Agence Régionale de Santé*) in France posts press releases with information on the policies the région or départements within the région will implement to mitigate the spread and impact of the COVID-19 outbreak.[58] The Wikipedia page on the 2020 coronavirus pandemic in France has collated information on the major policy measures taken in response to the COVID-19 pandemic.[59]

Starting February 29, 2020, France banned mass gatherings of more than 5,000 people nationwide, while some major sporting events were cancelled and a handful of schools closed to mitigate the spread of the virus. As more COVID-19 cases were confirmed during the following week, additional sporting events were canceled, more schools decided to close, and certain cities and *départements* limited mass gatherings to no more than 50 people, excluding shops, business, restaurants, bars, weddings, and funerals. Some régions closed early childhood establishments (e.g. nurseries, daycare centers) and prohibited visitors to elderly care facilities. On March 8, 2020, France banned mass gatherings of more than 1,000 people nationwide. Other schools, cities, and départements followed suit with additional school closures and limiting mass gatherings. On March 11, 2020, France prohibited all visits to elder care establishments. Starting March 16, 2020, France closed all schools nationwide.

We have coded various policies that cancel events and large gatherings as such: any cancellations of professional sporting and other specific pre-scheduled events as the policy variable "*event_cancel*." The "*no_gathering*" policy variable represents policy measures that banned all events or mass gatherings of a certain size, e.g. no gatherings of over 1,000 people. The "*social_distance*" policy variable includes measures preventing visits to elder care establishments, closures of public pools and tourist attractions, and teleworking plans for workers.

South Korea

We obtained data on South Korea's policy response to the COVID-19 pandemic from various news sources, as well as press releases from the Korean Centers for Disease Control and Prevention (KCDC), the Ministry of Foreign Affairs, and local governments' websites. The policy variables coded in the dataset are: "*business_closure,*" "*business_closure_opt,*" "*emergency_declaration,*" "*no_demonstration,*" "*religious_closure*," "*school_closure,*" "*social_distance_opt,*" "*travel_ban_intl_in_opt,*" "*travel_ban_intl_out_opt,*" and "*work_from_home_opt.*"

The KCDC announced on February 28, 2020 that health-related public facilities were recommended to be closed;[60] hence, the policy variable "*business_closure*" was coded as one from the announcement date. Even though it was technically a recommendation, we did not code this policy as an optional

[57] https://www.gouvernement.fr/info-coronavirus
[58] https://www.ars.sante.fr/
[59] https://fr.wikipedia.org/wiki/Pand%C3%A9mie_de_maladie_%C3%A0_coronavirus_de_2020_en_France
[60] http://www.mohw.go.kr/react/al/sal0301vw.jsp?PAR_MENU_ID=04&MENU_ID=0403&page=8&CONT_SEQ=353184

one because a majority of facility types listed in the press release (senior centers, job centers, children's centers, etc.) are under public administration, so these facilities likely would have followed recommendations. Indeed, some news articles have reported that all children's centers in Busan are closed[61] and over 3,600 facilities in Seoul.[62]

We have another business variable, "*business_closure_opt*", which applies to two provinces: Seoul and Gyeonggi-do. On March 11, 2020, the mayor of Seoul advised that popular commercial establishments such as karaoke places, clubs, and cyber cafes be closed.[63] Seven days later, the governor of Gyeonggi-do issued an executive order limiting the usage of commonly frequented commercial establishments and requiring a higher standard of cleanliness.[64] We coded this as an optional business closure given that the policy discourages usage of these facilities but did not explicitly order them to shut down.

Daegu and Gyeongsangbuk-do have been two of the regions hardest hit by COVID-19. The government of South Korea declared an emergency for those two areas on March 15, 2020.[65] We incorporated this information into the variable "*emergency_declaration*."

The variable "*no_demonstration*" reflects the efforts of some regions limiting any protests calling for slowing the spread of the outbreak. On February 24, 2020, Incheon stopped a protest in front of the Incheon Metropolitan City Hall.[66] Two days later, Seoul prohibited protests in downtown areas where massive demonstrations used to take place.[67]

Many province level COVID-19 policies have targeted religious gatherings at Shincheonji Church of Jesus, since its religious gatherings have been linked to the explosion in the number of cumulative confirmed cases. Provincial governments tried to shut down Shincheonji-related places of worship, and the related policy implementation is encoded in the variable "*religious_closure*." The regions which utilized this policy option are: Daegu,[68] Gyeongsangbuk-do,[69] Seoul,[70] Jeju,[71] Gyeonggi-do,[72]

---

[61] http://www.kookje.co.kr/news2011/asp/newsbody.asp?code=0300&key=20200313.33001005312
[62] http://www.hani.co.kr/arti/area/capital/929213.html
[63] http://news.seoul.go.kr/welfare/archives/512657
[64] https://gnews.gg.go.kr/briefing/brief_gongbo_view.do?BS_CODE=s017&number=43714&period_1=&period_2=&search=0&keyword=&subject_Code=BO01&page=1
[65] http://ncov.mohw.go.kr/tcmBoardView.do?brdId=3&brdGubun=31&dataGubun=&ncvContSeq=1241&contSeq=1241&board_id=311&gubun=BDC
[66] http://press.incheon.go.kr/citynet/jsp/sap/SAPNewsBizProcess.do?command=searchDetailSvp&sido=&matOfYmd=20200224&matSno=10&flag=&viFlag=in
[67] http://mediahub.seoul.go.kr/archives/1270860
[68] http://www.ctimes.co.kr/news/articleView.html?idxno=6843
[69] https://www.msn.com/ko-kr/news/national/%EA%B2%BD%EB%B6%81-%EC%8B%A0%EC%B2%9C%EC%A7%80-1612%EB%AA%85-%EC%A4%91-221%EB%AA%85-%ED%99%95%EC%A7%84%C2%B7%C2%B7%C2%B731%EB%B2%88%EC%9D%B4-156%EB%AA%85-%EC%98%AE%EA%B2%BC%EB%8B%A4/ar-BB10C1am

Jeollanam-do,[73] Gyeongsangnam-do,[74] Incheon,[75] Ulsan,[76] Busan,[77] Jeollabuk-do,[78] Chungcheongbuk-do,[79] Gwangju,[80] Chungcheongnam-do,[81] and Daejeon.[82]

The policy variable "*school_closure*" has been turned on for the entirety of the Korean time series dataset. This is because all schools were already on vacation during the beginning of the outbreak, and the government then postponed their start dates. At the time of writing, the Ministry of Education announced that schools would be kept closed until April 3, 2020.[83] Therefore, this policy variable is always equal to 1 in the dataset.

"*social_distance_opt*" has been turned on from February 29, 2020, when KCDC recommended social distancing as one of the main tools to deal with the outbreak. In their press release, they recommended that "people maintain personal hygiene and practice 'social distancing' until the beginning of March, an important point of this outbreak."[84] In the case of Daegu, the hardest-hit region in the country, we coded the variable as 1 starting from February 22, 2020, based on the statement, "It is recommended for residents in Daegu to minimize gathering events and outdoor activities."[85]

The first travel restriction for incoming travelers ("*travel_ban_intl_in_opt*") was implemented on January 28, 2020. It is worth noting that it was not a total prohibition of incoming visitors; rather, it means inbound travellers were subject to COVID-19 specific emergency measures. KCDC mentioned that starting on January 28, 2020 "any travellers depart[ing] from China [would] be a subject to

---

[70] http://www.c-herald.co.kr/news/articleView.html?idxno=2156
[71] http://www.jemin.com/news/articleView.html?idxno=646993
[72] http://www.kookje.co.kr/news2011/asp/newsbody.asp?code=0300&key=20200224.99099008869
[73] http://www.kwangju.co.kr/article.php?aid=1582729200690279004
[74] http://woman.chosun.com/mobile/news/view.asp?cate=C01&mcate=M1004&nNewsNumb=20200264476#_enliple
[75] https://www.gov.kr/portal/ntnadmNews/2102077
[76] https://www.yna.co.kr/view/RPR20200225010600353
[77] https://www.ajunews.com/view/20200311082759689
[78] http://www.segye.com/newsView/20200309517306
[79] http://www.cbnews.kr/news/articleView.html?idxno=113191
[80] http://www.bosa.co.kr/news/articleView.html?idxno=2122251
[81] http://www.dtnews24.com/news/articleView.html?idxno=572551
[82] https://blog.naver.com/PostView.nhn?blogId=storydaejeon&logNo=221834017110&redirect=Dlog&widgetTypeCall=true&directAccess=false
[83] https://www.moe.go.kr/boardCnts/view.do?boardID=294&boardSeq=80044&lev=0&searchType=null&statusYN=W&page=1&s=moe&m=020402&opType=N
[84] https://www.cdc.go.kr/board/board.es?mid=a30402000000&bid=0030&act=view&list_no=366406&tag=&nPage=2
[85] https://www.cdc.go.kr/board/board.es?mid=a30402000000&bid=0030&act=view&list_no=366299&tag=&nPage=3

strengthened screening and quarantine measures."[86] On February 12, 2020, KCDC broadened the list of countries subject to the stricter measures to include Hong Kong and Macau.[87] Subsequently, KCDC added Italy and Iran (on March 11, 2020)[88]; France, Germany, Spain, UK, and Netherlands (on March 15, 2020)[89] ; and any remaining European countries (March 15, 2020)[90] to their country list.

This restriction was not limited to inbound travellers. The government also issued advisories on countries where the number of infections had increased, which has been encoded as the variable "*travel_ban_intl_out_opt*." The first outbound travel alert due to COVID-19 was announced on January 28, 2020: The Ministry of Foreign Affairs (MOFA) issued a Level 2 (Yellow) alert for any travel to mainland China, Hong Kong, and Macau.[91] Later, MOFA added Italy on February 28, 2020,[92] Japan on March 9, 2020,[93] and all European countries on March 16, 2020.[94] It should be noted that the Level 2 alert does not enable the government to prohibit travel to these destinations, which is why the policy was coded as "optional."

There are four types of travel advisories distributed by the South Korean government: Level 1, Navy; Level 2, Yellow; Level 3, Red; and Level 4, Black.[95] Travel under the Level 4 alert is prohibited, and the government utilizes legal instruments to enforce the restriction. If people leave the country under the black alert, they will be subject to fines up to ten million KRW, or imprisonment up to a year. However, there is no enforcement instrument for the advisories up to Level 3. In that sense, we stated above that the banning policy does not mean prohibiting travel. Nevertheless, we coded the yellow alert as

86 https://www.cdc.go.kr/board/board.es?mid=a30402000000&bid=0030&act=view&list_no=365875&tag=&nPage=3

87 https://www.cdc.go.kr/board/board.es?mid=a30402000000&bid=0030&act=view&list_no=366154&tag=&nPage=1

88 https://www.cdc.go.kr/board/board.es?mid=a30402000000&bid=0030&act=view&list_no=366523&tag=&nPage=1

89 https://www.cdc.go.kr/board/board.es?mid=a30402000000&bid=0030&act=view&list_no=366537&tag=&nPage=1

90 https://www.cdc.go.kr/board/board.es?mid=a30402000000&bid=0030&act=view&list_no=366568&tag=&nPage=1

91 http://www.0404.go.kr/dev/newest_view.mofa?id=ATC0000000007598&pagenum=1&mst_id=MST00000000000040&ctnm=&div_cd=&st=title&stext=

92 http://www.0404.go.kr/dev/newest_view.mofa?id=ATC0000000007690&pagenum=1&mst_id=MST00000000000040&ctnm=&div_cd=&st=title&stext=

93 http://www.0404.go.kr/dev/newest_view.mofa?id=ATC0000000007709&pagenum=1&mst_id=MST00000000000040&ctnm=&div_cd=&st=title&stext=

94 http://www.0404.go.kr/dev/newest_view.mofa?id=ATC0000000007723&pagenum=1&mst_id=MST00000000000040&ctnm=&div_cd=&st=title&stext=

95 http://www.0404.go.kr/walking/walking_intro.jsp

the first travel ban in our dataset, since Level 2 alerts are issued relatively rarely, such as during a significant demonstration[96] or military coup.[97] As a result, we coded the Level 2 alert due to COVID-19 into the dataset for the policy analysis.

The policy variable "*work_from_home_optional*" indicates when KCDC began recommending that people work from home. On March 15, 2020, the KCDC press release stated: "Since contact with confirmed cases in an enclosed space increases the possibility of transmission, it is recommended to work at home or adjust desk locations so as to keep a certain distance among people in the office. More detailed guidelines for local governments and high-risk working environments will be distributed soon."[98]

Italy

We have obtained data on Italy's policy responses to the COVID-19 pandemic primarily from the English version of the COVID-19 dossier "Chronology of main steps and legal acts taken by the Italian Government for the containment of the COVID-19 epidemiological emergency"[99] written by the Department of Civil Protection (*Dipartimento della Protezione Civile*), most recently updated on March 12, 2020. This dossier details the majority of the municipal, regional, provincial, and national policies rolled out between the start of the pandemic to present-day. We have supplemented these policy events with news articles that detail which administrative areas were specifically impacted by the additional policies.

The first major policy rollout was on February 23, 2020, when 11 municipalities across two provinces in Northern Italy were placed on lockdown. These policies included closing schools, cancelling public and private events and gatherings, closing museums and other cultural institutions, closing non-essential commercial activities, and prohibiting the movement of people into or out of the municipalities.

The second major policy rollout was on March 1, 2020, when two provinces and three regions in Northern Italy were placed on partial lockdown. These policies also included closing schools, cancelling public and private events and gatherings, closing museums, closing non-essential commercial activities, as well as limiting the number of people at places of worship, restricting operating hours of bars and restaurants, and encouraging people to work remotely.

---

[96] http://www.0404.go.kr/dev/notice_view.mofa?id=ATC0000000007416

[97] http://www.0404.go.kr/dev/notice_view.mofa?id=8679&pagenum=1&st=title&stext=%EC%97%AC%ED%96%89%EA%B2%BD%EB%B3%B4

[98] https://www.cdc.go.kr/board/board.es?mid=a30402000000&bid=0030&act=view&list_no=366523&tag=&nPage=1

[99] http://www.protezionecivile.it/documents/20182/1227694/Summary+of+measures+taken+against+the+spread+of+C-19/c16459ad-4e52-4e90-90f3-c6a2b30c17eb

The third major policy roll-out was on March 5, 2020, when all schools across the country were closed.

The fourth major policy roll-out was on March 8, 2020 when the region of Lombardy and 13 provinces in Northern Italy were placed on lockdown. These policies included the cancellation of public and private events and gatherings, closing of museums, encouraging people to work remotely, limiting the number of people at places of worship, restricting opening hours of bars and restaurants, mandating quarantine of people who tested positive for COVID-19, prohibiting the movement of people into or out of the affected area, and restricting movement within the affected area to only work- or health-related purposes. Commercial activities were still allowed, as long as they maintained a safety distance of one meter apart per person within the establishment. All civil and religious ceremonies, including weddings and funeral ceremonies, were suspended. During this same policy roll-out, the rest of the country faced less stringent policies: cancelling of public and private events, closing of museums, and requiring restaurants and commercial establishments to maintain a safety distance of one meter apart per person within the establishment.

The fifth major policy roll-out was announced on March 9, 2020, and went into effect on March 10, 2020, when lockdown policies applied to Northern Italy were rolled out to the entire country. Lastly, on March 11, 2020, the lockdown was changed to also cover the closing of any non-essential businesses and further restricted people from leaving their home.

Iran

For Iran's policy response to the COVID-19 pandemic, we relied on news media reporting as the primary source of policy information (mostly due to translation restrictions). We also relied on two timelines of pandemic events in Iran to help guide the policy search.[100] [101]

The first major outbreak in Iran was connected to a major Shia pilgrimage in the city of Qom that brought Shiite pilgrims from Iran and throughout the Middle East, where they came to kiss the Fatima Masumeh shrine. It is possible that the disease was brought to Qom by a merchant traveling from Wuhan, China.[102] In addition, it is believed that the Iranian government knew of the COVID-19 outbreak prior to its February 21, 2020 parliamentary elections, but downplayed the risks associated with the disease as not to suppress voter turnout (given concerns that a low turnout would reflect poorly on its legitimacy).[103] The disease, initially centered in Qom and neighboring Tehran, spread rapidly throughout the country.

---

[100] https://www.thinkglobalhealth.org/article/updated-timeline-coronavirus
[101] https://en.wikipedia.org/wiki/2020_coronavirus_pandemic_in_Iran
[102] https://www.newyorker.com/news/our-columnists/how-iran-became-a-new-epicenter-of-the-coronavirus-outbreak
[103] https://www.newyorker.com/news/our-columnists/how-iran-became-a-new-epicenter-of-the-coronavirus-outbreak

As the number of cases grew, the Iranian government started to increase the stringency of its response. The first case was reported on February 19, 2020 (two individuals who both were reported to have died that day). The next day, school closures were announced in the province of Qom and travel in the region was discouraged. By February 22, 2020 the government closed schools in 14 provinces and closed down major gathering sites such as football matches and theaters. By March 5, 2020 schools were closed nationwide and government employees were required to work from home. Home isolation was implemented by the military on March 13, 2020, which the media described as "the near-curfew follows growing exasperation among MPs that calls for Iranian citizens to stay at home had been widely ignored, as people continued to travel before the Nowruz New Year holidays." [104]

United States

For the United States' policy response to the COVID-19 pandemic, we relied on a number of sources, including the U.S. Center for Disease Control (CDC), individual state health departments, as well as various press releases from county and city-level government or media outlets. The CDC has posted and continually updated a Community Mitigation Framework that encompasses both mandatory and recommended policies at a national level.[105] [106] This framework was interpreted by individual states as they each declared their own States of Emergency at various dates, and subsequently released their own community mitigation plans. Some of the first states to release such plans include Massachusetts, California, Florida, Washington, and New York.[107] Each respective Community Mitigation Framework included both mandatory and optional policies to prevent the COVID-19 spread. In addition to both national and federal level policies and recommendations, cities and counties have also taken on the role of providing guidance and implementing policies to mitigate the spread of COVID-19.

There have been a wide range in responses across states since the first case of COVID-19 was announced in Washington State on January 14, 2020. Upon this, the CDC began releasing recommendations to those at risk of being exposed to the virus. The initial recommendations included travel warnings and restricted travel to countries with confirmed cases and sustained COVID-19 spread. These travel restrictions grew to include inbound and outbound travel bans to a list of 26 countries, in both Europe and Asia.[108]

Other policies have included social distancing, which has been widely recommended or enforced at various levels of government. This method involves avoiding crowds, staying home, limiting or avoiding visiting vulnerable populations (such as long-term care facilities) and standing at least six feet away from others in public spaces.[109] [110] Some regions have implemented school closures at both

---

[104] https://www.theguardian.com/world/2020/mar/13/revolutionary-guards-enforce-coronavirus-controls-iran

[105] https://www.cdc.gov/coronavirus/2019-ncov/whats-new-all.html

[106] https://www.cdc.gov/coronavirus/2019-ncov/downloads/community-mitigation-strategy.pdf

[107] https://www.cdc.gov/coronavirus/2019-ncov/community/index.html

[108] https://www.cdc.gov/coronavirus/2019-ncov/travelers/after-travel-precautions.html

[109] https://www.nytimes.com/2020/03/16/smarter-living/coronavirus-social-distancing.html

the K-12 and higher education level. Business closures have also been recommended or enforced, such that employees should work from home, unless their work is considered essential to the greater public (e.g. health care, grocery stores). To support employees working remotely or staying home when sick, a number of states have also mandated paid sick leave for those who are affected by COVID-19. Free testing has also been implemented in certain states, so that anyone experiencing symptoms or has been exposed to the virus can now get tested for free.[111]

We coded various policies that cancel events and large gatherings as follows: the cancellation of large events, specifically the election postponement in Louisiana, is categorized as "*event_cancel*." The separate "*no_gathering*" policy variable represents policy measures that banned all events or mass gatherings of a certain size, i.e. no gatherings over a certain number of people (where this number has varied by region). The "*social_distance*" category includes measures that prevent visits to elderly care facilities, close public facilities such as libraries, and require workers to work remotely. The "*emergency_declaration*" encompasses the declarations of a state of emergency at the city, county, state, and federal level. This declaration allows the affected area to immediately marshal emergency funds and resources and activate emergency legislation, while also giving the public an indication of the gravity of the situation.

Population Data

In order to construct population weighted policy variables and to determine the susceptible fraction of the population for disease projections under the realized and the "no policy" counterfactual scenarios, we obtained the most recent estimates of population for each administrative unit included in our analysis. The sources of that population data are documented below.

China

City-level population data have been extracted from a compiled dataset of the 2010 Chinese City Statistical Yearbooks. We matched the city level population dataset to the city level COVID-19 epidemiology dataset. As the two datasets use slightly different administrative divisions, we only matched 295 cities that exist in both datasets, and grouped the remaining 39 cities in our compiled epidemiology dataset into "other" for prediction purposes. Cities grouped into "other" because of mismatches have a total population of 114,000,000, or 8.5% of the total population in China.

France

Département-level populations are obtained from the National Institute of Statistics and Economic database https://www.insee.fr/fr/statistiques/2012713#tableau-TCRD_004_tab1_departements.

We used the most up to date estimation of the population in France as of January 2020.

[110] https://www.cdc.gov/coronavirus/2019-ncov/community/large-events/mass-gatherings-ready-for-covid-19.html

[111] https://appropriations.house.gov/sites/democrats.appropriations.house.gov/files/Families%20First%20summary.pdf

South Korea

We downloaded the number of population by provinces from a webpage administered by the Korean Statistical Information Service (KOSIS).[112] The government agency recently updated the population information of February, 2020, which we used for our analysis.

Italy

Region and province level population data come from the Italian National Institute of Statistics (Istat), estimating total population on January 1, 2019. The datasets for all Italian regions and provinces are scraped from Istat's website in *get_adm_info.ipynb*.

Iran

Province level population data for Iran comes from the 2016 Census, as listed on the *City Population* website.[113] It is scraped in *get_adm_info.ipynb*.

United States

State and county level population data come from the 2017 American Community Surveys dataset, and is downloaded via the *census* Python package[114] in *get_adm_info.ipynb.*

[112] http://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1B040A3&vw_cd=MT_ZTITLE&list_id=A6&seqNo=&lang_mode=ko&language=kor&obj_var_id=&itm_id=&conn_path=MT_ZTITLE

[113] https://www.citypopulation.de/en/iran/admin/

[114] https://github.com/datamade/census